Correction automatique pour les codes linéaires

Malika More
(malika.more@iut.u-clermont1.fr)
Alex Esbelin
(Alex.Esbelin@math.u-bpclermont.fr)

IREM de Clermont-Ferrand

Formation ISN

30 Juin 2011

Plan du cours

Introduction

Question

- Un message reconnu erroné est corrigé en le remplaçant par le mot de code le plus proche
- Comment calculer ce mot de code ?

Vecteur de correction

- Message reçu : M
- Correction

$$C = M \oplus V$$

  - Modification de certains bits de M
  - Au moyen d'un *vecteur de correction* V
  - Pour obtenir un message corrigé C
- La correction est probablement correcte lorsque le message corrigé C est le mot de code le plus proche du message reçu M

Vecteur de correction et matrice de contrôle

- Correction : $C = M \oplus V$
- Si la correction a atteint son but, alors C est un mot de code, donc le syndrome de C est nul
  - $HC = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$
- Par conséquent, le vecteur de correction V doit avoir le même syndrome que le message reçu M

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = HC = H(M \oplus V) = HM \oplus HV \Longrightarrow HV = HM = S$$

Vecteur de correction et probabilités

- Correction : $C = M \oplus V$
- Si la correction a atteint son but, alors C est le mot de code le plus proche de M
  - La correction doit changer le moins de bits possible de M
- Autrement dit, le vecteur de correction V comporte le moins de 1 possible
- Par conséquent, le vecteur de correction V doit avoir un poids le plus faible possible

Correction automatique

Principe

- On prend comme vecteur de correction V un vecteur qui a
  - Même syndrome que le message reçu M
  - Poids le plus faible possible
- On corrige en calculant

$$C = M \oplus V$$

- Le mot de code C obtenu est le mot de code le plus proche du message reçu M

Calcul du vecteur de correction

- Pour pouvoir corriger un message reçu M erroné, il faut trouver un vecteur de correction V qui a
  - Même syndrome que le message reçu M
  - Poids le plus faible possible
- S'il faut chaque fois :
  - Calculer les syndromes de tous les vecteurs
  - Trier tous les vecteurs qui ont le même syndrome que M
  - Parmi ces vecteurs, sélectionner celui qui a le poids le plus faible
- Ça peut être très long...

Tableau standard

- Pour gagner du temps, on précalcule UNE SEULE FOIS, pour chaque syndrome S, un vecteur V de poids le plus faible possible qui a pour syndrome S
  - Ce vecteur V servira de vecteur de correction pour TOUS les message reçus M de syndrome S
  - On obtient le *tableau standard* du code
- Aprés calcul du syndrome S du message reçu M, le choix du vecteur de correction V s'effectue par simple lecture du tableau standard

Exemple

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

mot d'info		mot de code
00	$\longrightarrow$	0000
01	$\longrightarrow$	0111
10	$\longrightarrow$	1010
11	$\longrightarrow$	1101

- Longueur des mots d'info $m = 2$ bits
  - $\rightsquigarrow$ Nombre de mots de code = Nombre de mots d'info = $2^m = 4$
- Longueur des mots de code $n = 4$ bits
  - $\rightsquigarrow$ Nombre messages $= 2^n = 16$
- Longueur des syndromes $r = n - m = 2$ bits
  - $\rightsquigarrow$ Nombre de syndromes $= 2^r = 4$

Exemple : Vecteurs classés par syndrome

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Syndrome	00	01	10	11
Vecteurs	0000 1010 0111 1101	0001 1011 0110 1100	1000 0010 1111 0101	0100 1110 0011 1001
Vecteur(s) de plus faible poids	0000	0001	1000 ou 0010	0100

Exemple : Vecteurs classés par syndrome

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Syndrome	00	01	10	11
Vecteurs	0000 1010 0111 1101	0001 1011 0110 1100	1000 0010 1111 0101	0100 1110 0011 1001
Vecteur(s) de plus faible poids	0000	0001	1000 ou 0010	0100

Exemple : Vecteurs classés par syndrome

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Syndrome	00	01	10	11
Vecteurs	0000 1010 0111 1101	0001 1011 0110 1100	1000 0010 1111 0101	0100 1110 0011 1001
Vecteur(s) de plus faible poids	0000	0001	1000 ou 0010	0100

Exemple : Choix du vecteur de correction

Syndrome	00	01	10	11
Vecteur(s) de plus faible poids	0000	0001	1000 ou 0010	0100
Vecteur de correction	0000	0001	NC_1	0100

- Syndrome S avec plusieurs vecteurs de plus faible poids k
  - Tous la même probabilité d'être le vecteur d'erreur
- Message reçu M avec syndrome S
  - Plusieurs mots de code à distance k de M
  - Impossible de corriger M par le mot de code le plus proche
- Dans le tableau standard, on écrit NC_k
  - Le syndrome S correspond à des messages *Non Corrigeables* avec k erreurs

Exemple : Choix du vecteur de correction

Tableau Standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

- Syndrome S avec plusieurs vecteurs de plus faible poids k
  - Tous la même probabilité d'être le vecteur d'erreur
- Message reçu M avec syndrome S
  - Plusieurs mots de code à distance k de M
  - Impossible de corriger M par le mot de code le plus proche
- Dans le tableau standard, on écrit NC_k
  - Le syndrome S correspond à des messages *Non Corrigeables* avec k erreurs

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1100$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1100$
- Calcul du syndrome $S = HM = 01$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1100$
- Calcul du syndrome $S = HM = 01$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1100$
- Calcul du syndrome $S = HM = 01$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

- Correction

$$C = M \oplus V = 1100 \oplus 0001 = 1101$$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1010$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1010$
- Calcul du syndrome $S = HM = 00$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1010$
- Calcul du syndrome $S = HM = 00$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Exemple : Utilisation du tableau standard

- Message reçu : $M = 1010$
- Calcul du syndrome $S = HM = 00$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

- Correction

$$C = M \oplus V = 1010 \oplus 0000 = 1010$$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 0101$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 0101$
- Calcul du syndrome $S = HM = 10$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Exemple : Utilisation du tableau standard

- Message reçu : $M = 0101$
- Calcul du syndrome $S = HM = 10$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Exemple : Utilisation du tableau standard

- Message reçu : $M = 0101$
- Calcul du syndrome $S = HM = 10$

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

- Lecture du tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

- Correction

$M = 0101$ message erroné non corrigeable (1 erreur)

Résumé : Les outils des codes linéaires

Codage : Matrice génératrice

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Détection : Matrice de contrôle

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Correction : Tableau standard

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Méthode théorique :
  - Calculer le syndrome de tous les 2^n vecteurs possibles
  - Classer les vecteurs par syndrome
  - Sélectionner chaque fois le(s) vecteur(s) de plus faible poids
- Méthode pratique :
  - Calculer les syndromes des vecteurs par *poids croissant*
  - Quand un nouveau syndrome apparaît, noter le vecteur correspondant
  - S'arrêter dès que le tableau est terminé (attention à NC_k)

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Construction du tableau standard

- Ligne 1 : Liste des 2^r syndromes
- Ligne 2 : Liste des vecteurs de correction correspondants
  - Vecteur nul : Syndrome nul
  - Vecteurs de poids 1 : Syndromes qui sont des colonnes de H
  - En cas d'ambiguité : Non Corrigeable d'ordre k
  - Poids 2,3,... si nécessaire

Syndrome	00	01	10	11
Vecteur de correction	0000	0001	NC_1	0100

Notion essentielle

Mesurer l'efficacité du code

- Messages erronés avec

  Strictement moins de d erreurs : TOUS DÉTECTÉS

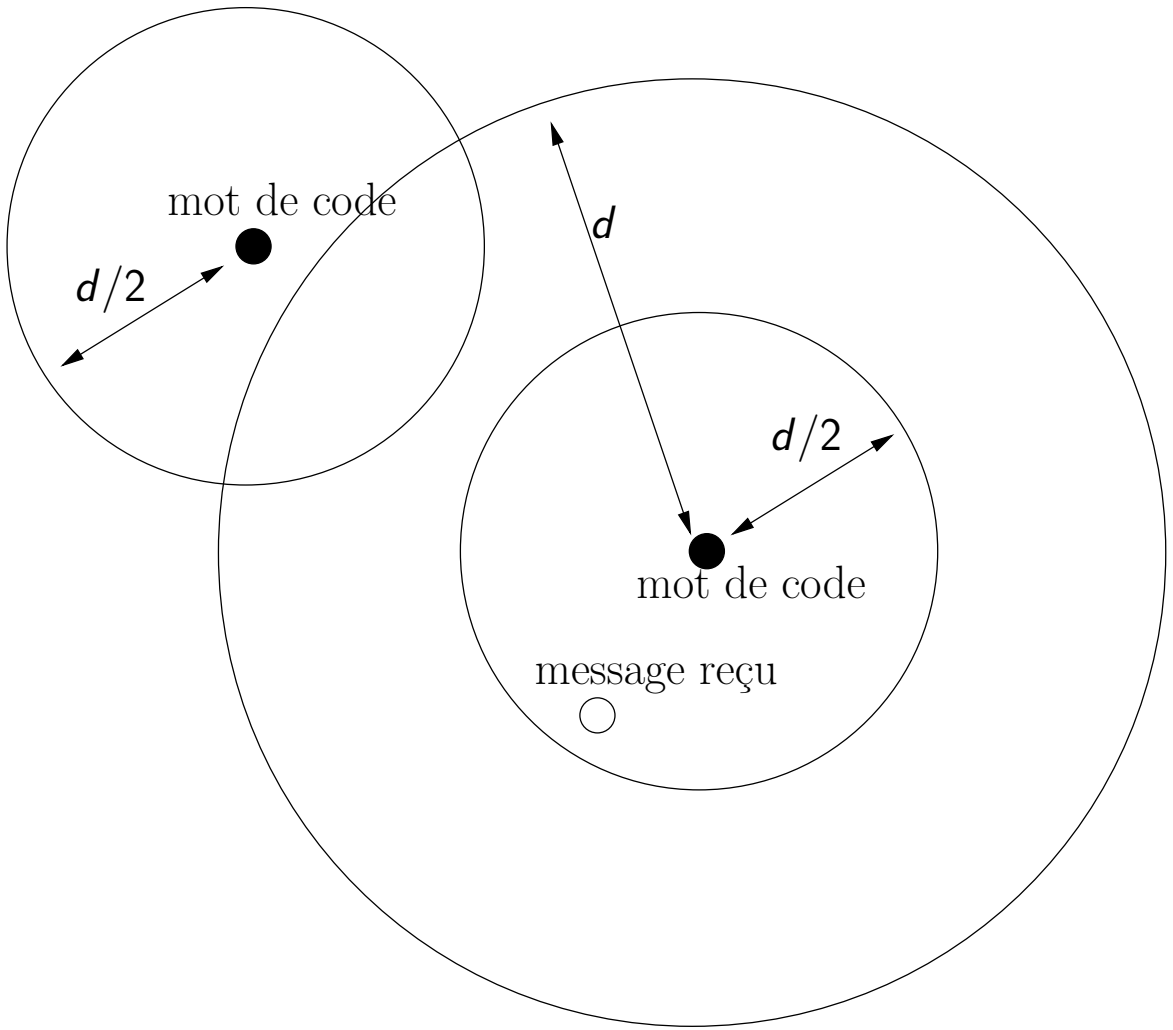
  Les autres : certains pas détectés

- Messages erronés avec

  Strictement moins de $d/2$ erreurs : TOUS BIEN CORRIGÉS

  Les autres : certains mal corrigés

Résumé



Distance minimale du code

- La distance minimale du code d est la plus petite distance de Hamming entre deux mots de code différents
- Pour calculer cette distance minimale pour un code quelconque, il faudrait
  - prendre toutes les paires de mots de code différents
  - calculer à chaque fois leur distance de Hamming
  - garder la plus petite valeur obtenue

  Pour un code $\mathcal{C}_{n,m}$, il y a 2^m mots de code... C'est long
- Dans le cas des codes *linéaires*, on peut le faire plus simplement

Mot de code non nul de plus faible poids

- Soient C_1 et C_2 deux mots de code avec $C_1 \neq C_2$
- La distance de Hamming entre C_1 et C_2 est $w(C_1 \oplus C_2)$
- Pour un code *linéaire*, $C_2 \oplus C_1 = C$ est aussi un mot de code (non nul car $C_1 \neq C_2$)
- Donc la distance de Hamming entre C_1 et C_2 est $w(C)$
- Conclusion :
  - Pour chercher la plus petite distance de Hamming entre deux mots de code, il suffit de chercher le plus petit poids possible pour un mot de code non nul

Distance Minimale d'un code linéaire

Propriété

La distance minimale d'un code *linéaire* est égale au poids d'un mot de code non nul de plus faible poids

- Mais, si on doit faire la liste de tous les mots de code non nuls pour calculer leur poids, ça peut encore être long...

Exemple

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

mot d'info		mot de code
00	$\longrightarrow$	0000
01	$\longrightarrow$	0111
10	$\longrightarrow$	1010
11	$\longrightarrow$	1101

La distance minimale d'un code *quelconque* est égale à la plus petite distance de Hamming entre deux mots de code différents

$w(1010 \oplus 0111) = 3 \quad w(1010 \oplus 0000) = 2$
$w(1010 \oplus 1101) = 3 \quad w(0111 \oplus 0000) = 3$
$w(0111 \oplus 1101) = 2 \quad w(0000 \oplus 1101) = 3$

$$d = 2$$

La distance minimale d'un code *linéaire* est égale au poids d'un mot de code non nul de plus faible poids

$w(1010) = 2 \quad w(0111) = 3 \quad w(1101) = 3$

$$d = 2$$

Colonnes de la matrice de contrôle

- Les colonnes de H sont les syndromes des messages de poids 1

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

- Plus précisément
  - La colonne C_i est le syndrome du message $m_i = 0 \dots 010 \dots 0$ avec un 1 en position i
- Les mots de code ont le syndrome nul $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ et eux seuls

Étude des colonnes de H

- H contient la colonne nulle : le message de poids 1 correspondant est un mot de code

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Conséquence : $d = 1$
  - Ce code ne permet ni de détecter ni de corriger les erreurs

Étude des colonnes de H

- H ne contient pas la colonne nulle : il n'y a pas de mot de code de poids 1

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

- Conséquence : $d \geq 2$
  - Ce code permet de détecter au moins une erreur, mais pas forcément de corriger

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H contient deux colonnes égales : le message de poids 2 correspondant est un mot de code

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Conséquence : $d = 2$
  - Ce code permet de détecter une erreur, mais ne permet pas de corriger

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales : il n'y a pas de mot de code de poids 2

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

- Conséquence : $d \geq 3$
  - Ce code permet de détecter au moins deux erreurs, et d'en corriger au moins une

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales
- H contient trois colonnes de somme $\oplus$ nulle : le message de poids 3 correspondant est un mot de code

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Conséquence : $d = 3$
  - Ce code permet de détecter deux erreurs, et d'en corriger une

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales
- H ne contient pas trois colonnes de somme $\oplus$ nulle : il n'y a pas de mot de code de poids 3

- Conséquence : $d \geq 4$
  - Ce code permet de détecter au moins trois erreurs, et d'en corriger au moins une

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales
- H ne contient pas trois colonnes de somme $\oplus$ nulle
- H contient quatre colonnes de somme $\oplus$ nulle : le message de poids 4 correspondant est un mot de code

- Conséquence : $d = 4$
  - Ce code permet de détecter trois erreurs, et d'en corriger une

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales
- H ne contient pas trois colonnes de somme $\oplus$ nulle
- H ne contient pas quatre colonnes de somme $\oplus$ nulle : il n'y a pas de mot de code de poids 4

- Conséquence : $d \geq 5$
  - Ce code permet de détecter au moins quatre erreurs, et d'en corriger au moins deux

Étude des colonnes de H

- H ne contient pas la colonne nulle
- H ne contient pas deux colonnes égales
- H ne contient pas trois colonnes de somme $\oplus$ nulle
- H ne contient pas quatre colonnes de somme $\oplus$ nulle
- H contient cinq colonnes de somme $\oplus$ nulle : le message de poids 5 correspondant est un mot de code

- Conséquence : $d = 5$
  - Ce code permet de détecter quatre erreurs, et d'en corriger deux

Calcul de la distance minimale du code

Principe

La distance minimale du code d est le plus petit nombre de colonnes de la matrice de contrôle H qu'il faut additionner $\oplus$ pour obtenir la colonne nulle.

FIN